Request to Archive
With The National Centers for Environmental Information
For WRF high resolution Hurricane Sandy simulation
Provided by Cray Inc.

2015-12-18

This information will be used by NCEI to conduct an appraisal and make a decision on the request.

1. Who is the primary point of contact for this request?

Peter Johnsen
Cray Inc.
Application Engineer
651-605-9173
pjj@cray.com

2. Name the organization or group responsible for creating the dataset.

Cray Inc. / NCSA / NSF / NCAR

3. Provide an overview summarizing the scope of data you want to archive. Describe the outputs, data variables, including their measurement resolution and coverage.

High resolution output from a 4 day WRF (NCAR Weather Research and Forecast model) simulation of Hurricane Sandy.
This simulation was run on NCSA Cray XE6 supercomputer during early 2013. Datasets are archived on NCSA nearline storage but we would like to make it publicly available for various research purposes.

Grid size: 5320x5000 horizontal with 150 vertical levels
Grid resolution: 500 meters
Simulation days: 4
Start date: 12Z October 26, 2012
End date: 12Z October 30, 2012
Format: NetCDF 3
Frequency: every 30 simulation minutes
Total files: 193
Size per file: 224 GBytes
Total size: 43.2 TBytes
WRF Version: 3.3.1

4. What is the time period covered by the dataset? (YYYY-MM-DD, YYYY-MM or YYYY)

From 2012-10-26 to 2012-10-30

5. Edition or version number(s) of the dataset:

N/A

6. Approximate date when the dataset was or will be released to the public:
2015-12-31

7. Who are the expected users of the archived data? How will the archived data be used?
Various meteorological researchers including NOAA National Hurricane Center (NHC).
Expect use to span the next several years.

8. Has the dataset undergone user evaluation and/or an independent review process? Did NCEI participate in design reviews?
Analyzed by NCAR researchers who produced post-processed output and visualizations.

9. Describe the dataset's relationship to other archived datasets, such as earlier versions or related source data. If this is a new version, how does it improve upon the previous version(s)?
Related to high resolution Hurricane Sandy simulation published at SuperComputing 2013 Conference in November 2013.
See web page/other links below.

10. List the input datasets and ancillary information used to produce the data.
WRF simulation initialized with NOAA/NCEP GFS global model output.

11. List web pages and other links that provide information on the data.
http://www.climatecentral.org/blogs/the-most-detailed-visualizations-of-hurricane-sandy-revealed-16563

http://www.hpcwire.com/2013/11/14/behind-blue-waters-hurricane-sandy-simulation/

http://sc13.supercomputing.org/schedule/event_detail.php?evid=pap255

12. List the kinds of documents, metadata and code that are available for archiving. For example, data format specifications, user guides, algorithm documentation, metadata compliant with a standard such as ISO 19115, source code, platform/instrument metadata, data/process flow diagrams, etc.
1. NetCDF 3

13. Indicate the data file format(s).
1. netCDF-3

14. Are the data files compressed?
No

15. Provide details on how the files are named and how they are organized (e.g., file_name_pattern_YYYYMM.tar in monthly aggregations).
Sub-directories containing 6 hours of simulation data each, every 30 simulation minutes.

Sample naming convention:

out_12zOct26/auxhist1_d01_2012-10-26_12_00_00

16. Explain how to access sample data files and/or a file listing for previewing. If it is not available now, when will it be available?

Will include text file containing file names and NetCDF ncdump header information.

17. What is the total data volume to be submitted?

Historic Data: all historic data or data submitted as a completed collection.

Total Data Volume: 43.2TB

Number of Data Files: 193

18. Are later updates, revisions or replacement files anticipated? If so, explain the conditions for submitting these additional data to the archive.

No additional updates, revisions or replacement data are anticipated.

19. Describe the server that will connect to the ingest server at NCEI for submitting the data.

Physical Location: NCSA Supercomputer Center, Champaign, Illinois USA

System Name: BlueWaters

System Owner: NCSA/NSF

Additional Information:

20. What are the possible methods for submitting the data to NCEI? Select all that apply.

Globus Online (GridFTP)

21. Identify how you would like NCEI to distribute the data. Web access support depends on the resources available for the dataset.

1. User interface to order and stage data for download
2. Direct download links

22. Will there be any distribution, usage, or other restrictions that apply to the data in the archive?

Constraint Type	Description
Use	Acknowledgments required for use and publication.

23. Discuss the rationale for archiving the dataset and the anticipated benefits. Mention any risks associated with not archiving the dataset at NCEI.

As far as we know this is the most complete and detailed simulation datasets available for an important weather event. No risks to NCEI expected.

24. Are the data archived at another facility or are there plans to do so? Please explain.

On NCSA nearline storage, not publicly available.

25. Is there an existing agreement or requirement driving this request to archive? Have you already contacted someone at NCEI?

No

26. Do you have a data management plan for your data?

No

27. Have funds been allocated to archive the data at NCEI?

No

28. Identify the affiliated research project, its sponsor, and any project/grant ID as applicable.
National Science Foundation (award number OCI 07-25070) and the state of Illinois. Blue Waters is a joint effort of the University of Illinois at Urbana-Champaign and its National Center for Supercomputing Applications.

Cray Inc.

29. Is there a desired deadline for NCEI to archive and provide access to the data?
No deadlines for archive or access.

30. Add any other pertinent information for this request.
None